

APPELLANT'S BRIEF UNDER 37 CFR §41.37		Docket No. C0012/7000
Applicant:	Raju C. Bopardikar, Jack J. Stiffler, Jacob Y. Bast, Gary A. Cardone, David E. Kaufman, Stuart P. MacEachern, Bruce D. McLeod, James M. Nolan, Jr., Zdenek Radouch and James A. Wentworth, III	
Serial No:	09/608,521	
Filed:	June 30, 2000	
For:	METHOD AND APPARATUS FOR IMPLEMENTING HIGH-PERFORMANCE, SCALEABLE DATA PROCESSING AND STORAGE SYSTEMS	
Examiner:	A. M. Mirza	
Art Unit:	2145	

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

This brief is in furtherance of the Notice of Appeal, filed in this case on November 16, 2011.

The fees required under §41.20(b)(2), and any required petition for extension of time for filing this brief and fees therefor, are dealt with in the accompanying TRANSMITTAL OF APPEAL BRIEF.

A single copy of this brief is transmitted (37 C.F.R. §41.37(a)) and contains these items under the following headings, and in the order set forth below (37 C.F.R. §41.37(c)(1)):

- I REAL PARTY IN INTEREST
- II RELATED APPEALS AND INTERFERENCES
- III STATUS OF CLAIMS
- IV STATUS OF AMENDMENTS
- V SUMMARY OF CLAIMED SUBJECT MATTER
- VI GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL
- VII ARGUMENT
- VIII APPENDIX OF CLAIMS INVOLVED IN THE APPEAL
- IX EVIDENCE APPENDIX
- X RELATED PROCEEDINGS APPENDIX

I REAL PARTY IN INTEREST (37 C.F.R. §41.37(c)(1)(i))

The real party in interest in this appeal is EMC Corporation, 176 South Street, Hopkinton, MA 01748.

II RELATED APPEALS AND INTERFERENCES (37 C.F.R. §41.37(c)(1)(ii))

There are no appeals or interferences that may directly affect, or be directly affected by, or have a bearing on the Board's decision in the pending appeal.

III STATUS OF CLAIMS (37 C.F.R. §41.37(c)(1)(iii))

A. TOTAL NUMBER OF CLAIMS IN APPLICATION

The claims in the application are: 1-86.

B. STATUS OF ALL THE CLAIMS IN APPLICATION

1. Claims pending: 1, 3-11, 44 and 46-54.
2. Claims canceled: 2 and 45.
3. Claims withdrawn from consideration, but not canceled: 12-43 and 55-86.
4. Claims allowed: None.
5. Claims rejected: 1, 3-6, 9-11, 44, 46-49 and 52-54.
6. Claims objected to 7-8 and 50-51.

C. CLAIMS ON APPEAL

The claims on appeal are: 1, 3-11, 44, 46-54.

IV STATUS OF AMENDMENTS (37 C.F.R. §41.37(c)(1)(iv))

No amendments have been filed.

V SUMMARY OF CLAIMED SUBJECT MATTER (37 C.F.R. §41.37(c)(1)(v))

The present invention relates to a method and apparatus for providing storage services to clients from a “pool” of storage devices (see page 4, line 21 - page 5, line 13 and Figure 1). The clients (102, 104) are connected to access interfaces (106, 108, 122) which are, in turn, connected to the pool of storage devices (118, 120, 122, 124) by a switch fabric (114, 116). The access interfaces communicate with the devices in the storage device pool and, in response to each storage request and based on the

workload of each device in the storage device pool, select a subset of the storage devices to use for any given transaction and distribute the workload (page 5, lines 16-22). Thus, the data received from a client in each storage request is transferred to the selected subset of storage devices based on the workload of the devices rather than on a physical device address (page 5, line 23 – page 6, line 6). The result of this operation is that data may be stored in a storage device other than the storage device from which that data was retrieved (page 7, line 22 – page 8, line 11).

The independent claims are supported by the specification as follows:

Claim 1 - an access interface module (page 5, lines 9-14)

which receives from the client data storage requests, each including a data object to be stored and a data object identifier that identifies that data object (page 10, lines 11-25) and,

in response to each storage request and based on a workload and on relative demands placed on subsets of the plurality of storage devices instead of a physical location in the plurality of devices, dynamically selects a subset of the plurality of storage devices in which the data is stored (page 5, lines 16-22 and page 14, lines 12-18, page 18, lines 12-24)

so that data corresponding to the same data object identifier can be transferred to different physical storage device locations from request to request in order to dynamically distribute the workload across the plurality of storage devices (page 14, line 19-page 15, line 2); and

a switch fabric for temporarily connecting the access interface module to the selected subset of the plurality of storage devices so that the

data can be transferred to the selected subset of storage devices
(page 4, line 25 – page 5, line 8).

Claim 44

step (a) - page 10, lines 11-25

step (b) - page 5, lines 16-22 and page 14, lines 12-18; page 14, line 19-
page 15, line 2

step (c) - page 4, line 25 – page 5, line 8

The dependent claims are supported by the specification as follows:

Claim 3 – page 4, line 25 – page 5, line 1, Figure 9, 952 and 942.

Claim 4 - page 5, lines 1-8.

Claim 5 – page 10, lines 4-10 Figure 3, 302, 304.

Claim 6 – page 22, lines 21-28, Figure 9, 952 and 942.

Claim 7 – page 12, lines 11-15 Figure 2, 218, 220, 222, 224.

Claim 8 – page 12, lines 15-17, Figure 2, 206, 216, 218.

Claim 9 – page 12, lines 17-18.

Claim 10 – page 9, lines 14-18, Figure 3, 322.

Claim 11 – page 4, lines 23-24, Figure 1 106, 108, 110, 112.

Claim 46 – page 4, line 25 – page 5, line 1, Figure 9, 952 and 942.

Claim 47 - page 5, lines 1-8.

Claim 48 – page 10, lines 4-10 Figure 3, 302, 304.

Claim 49 – page 22, lines 21-28, Figure 9, 952 and 942.

Claim 50 – page 12, lines 11-15 Figure 2, 218, 220, 222, 224.

Claim 51 – page 12, lines 15-17, Figure 2, 206, 216, 218.

Claim 52 – page 12, lines 17-18.

Claim 53 – page 9, lines 14-18, Figure 3, 322.

Claim 54 – page 4, lines 23-24, Figure 1 106, 108, 110, 112.

VI GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL (37 C.F.R. 41.37 (c)(1)(vi))

- A. Whether claims 1, 3-11, 44 and 46-54 are patentable under 35 U.S.C. 103(a) over U.S. Patent No. 6,237,063 (Bachmat) in view of U.S. Patent No. 5,909,686 (Muller.)

VII ARGUMENT (37 C.F.R. §41.37(c)(1)(vii))

- A. Prima facie obviousness has not been established because the combination of Bachmat and Muller does not teach or suggest the structure recited in claims 1, 3-11, 44 and 46-54.**

Obviousness is a legal conclusion based on factual evidence. Graham v. John Deere Co. 383 US 1, 148 USPQ 459 (1966). To establish the *prima facie* obviousness of a claimed invention, all of the claim limitations must be taught or suggested by the

prior art. In re Wada and Murphy, Appeal 2007-3733, citing In re Ochiai, 71 F.3d 1565, 1572 (Fed. Cir. 1995) and In re Royka, 490 F.2d 981, 180 USPQ 580 (CCPA 1970).

The Bachmat reference discloses load balancing of activities on physical disk storage devices by monitoring reading and writing operations to blocks of contiguous storage locations on the physical disk storage devices. A list of exchangeable pairs of blocks is developed based on size and function. Statistics accumulated over an interval are then used to obtain access activity values for each block and each physical disk drive. A statistical analysis leads to a selection of one block pair. After testing to determine any adverse effect of making that change, the exchange is made to more evenly distribute the loading on individual physical disk storage devices.

Although the Bachmat reference may be directed towards a similar goal as the claimed invention, Bachmat functions differently from what is recited in the claims. As disclosed in Bachmat, the movement of the data to accomplish load balancing is performed at discrete intervals after data read and write operations have occurred. This type of operation is necessary in order to gather and compile data on reading and writing. See, for example, Bachmat, column 3, lines 21-28:

“In accordance with this invention load balancing will occur at some arbitrary time interval, typically after an interval of one or more days. To that point in time various reading and writing statistics are accumulated to different blocks of contiguous storage locations. The compiled data is then used to select two blocks as candidates for an exchange and thereafter exchanging the data in the selected logical blocks.

Further, in Bachmat, blocks of data, such as logical volumes, are swapped in order to accomplish load balancing. See, for example, Bachmat, column 3, lines 29-45:

“... load balancing activity occurs with respect to a plurality of physical disk storage devices in a data processing system wherein at least two of the physical disk storage devices are divided into a plurality of logical volumes ... Disk access statistics are compiled for all the logical volumes over a time interval. These statistics are used to select first and second logical volumes on different physical disk storage devices to be exchanged based upon the compiled disk access statistics. Once the selection is made, the (a) data [sic] in the selected first and second logical volumes are

exchanged by transfer through the first and second buffer volumes.”
(emphasis added)

In contrast, the claimed invention accomplishes load balancing in response to each storage request as recited in claim 1, lines 3-11:

“... an access interface module which receives from the client data storage requests, each including a data object to be stored and a data object identifier that identifies that data object and, in response to each storage request and based on a workload and on relative demands placed on subsets of the plurality of storage devices instead of a physical location in the plurality of storage devices, dynamically selects a subset of the plurality of storage devices in which the data is stored so that data corresponding to the same data object identifier can be transferred to different physical storage device locations from request to request in order to dynamically distribute the workload across the plurality of storage devices ...” (emphasis added)

It is clear that the Bachmat load balancing operation does not disclose the features of the claimed invention. Bachmat is concerned with collecting data over time and then performing statistical calculations on the collected data; only then is data moved. The claims do not recite studying data access patterns as in Bachmat, but rather recite the feature that “in response to each storage request ... dynamically selects a subset of the plurality of storage devices in which the data is stored ... in order to dynamically distribute the workload across the plurality of storage devices”. This operation involves, among other things, a feature that “preallocates storage ... in order to evenly distribute a workload across the plurality of storage devices” as recited, for example, in claim 7 (see discussion below). Bachmat neither teaches nor suggests this type of operation because it is not necessary for the Bachmat system.

The Muller reference discloses a store and forward data packet switch in which data packets received at input ports are temporarily stored and then selectively forwarded to switch output ports based on information in a forwarding database. The bulk of the Muller disclosure is directed to hardware that allows a CPU to efficiently manage the information in the forwarding database as the switch operates. However,

the Muller reference does not disclose load balancing of physical storage devices or any method for accomplishing this operation.

Accordingly, Bachmat and Muller do not teach a system that “in response to each storage request ... dynamically selects a subset of the plurality of storage devices in which the data is stored” as recited in claims 1 and 44. Accordingly, claims 1 and 44 patentably distinguish over the cited reference combination.

The examiner cites Bachmat column 12, line 63, to column 13, line 8, as disclosing these limitations. However, at this location Bachmat discusses the sizes of blocks of data that are swapped in order to accomplish load balancing. Bachmat makes it clear that a “block” constitutes a group of contiguous storage locations, not a single storage location:

“Each block, whether a logical volume, sub-volume or other grouping, constitutes a block of contiguous storage locations of a predetermined size. Conversely and consequently, a block then can be a single logical volume, sub-volume or other grouping”

Further, Bachmat does not disclose at this or any other location performing load balancing in response to each data storage request.

The examiner further states that “One ordinary skill in the art at the time of the invention knows that ‘routing information based on the loading characteristic of the storage access request received at the switch inputs in order to balance loading of the storage access requests upon the outputs of the switch’ is prime example of load balancing, Where the routing information is based on the loading characteristic of the meaning workload of the memory” However, the quoted passage does not appear in either of the Bachmat or Muller references.

Claims 3-6, 9-11, 46-49 and 52-54 are dependent, either directly or indirectly, on claims 1 and 44, respectively, and incorporate the limitations thereof. Therefore, they also distinguish over the cited reference combination in the same manner as claims 1 and 44. In addition, these claims recite further elements not disclosed or suggested by the cited reference combination.

Claims 3-6 and 46-49 recite additional features. Claims 3 and 46 cite the features that “wherein the switch fabric comprises a control switch fabric for transferring control information and a separate data switch fabric for transferring data” Claims 4 and 47 recite the features that “the control switch fabric is optimized for transferring control information and the data switch fabric is optimized for transferring data.” Claims 5 and 48 recited the features that “the request for storage includes control information and data and wherein the access interface module separates the control information and the data and transfers the data to the selected subset of storage devices over the data switch fabric.” Claims 6 and 49 recite the features of “the data switch fabric comprises a non-blocking crossbar switch for data transfer and the control switch fabric comprises an Ethernet switch for control information transfer.”

With regard to claims 3 and 46, the elements recited in claims 3 and 46, “a control switch ... and a separate data switch fabric” are not disclosed in Bachmat. Bachmat does not utilize a switch to control access to the storage devices and instead relies on bus access logic to control this access. See Bachmat, column 4, line 65 to column 5, line 6. Muller also discloses no switch analogous to the recitation of these claims. The examiner refers to Muller column 4, lines 37-43 as disclosing this limitation. However only a single block switch element 210 is disclosed at this location. The contents of this block are further disclosed in Muller Figure 3 and column 7, line 8 to column 10 line 8. There is no teaching or suggestion of in these locations of two switch elements as recited in claims 3 and 46.

Claims 4 and 47 recite that “the control switch fabric is optimized for transferring control information and the data switch fabric is optimized for transferring data”. The examiner cites Muller, column 5, lines 37-48 as disclosing this limitation. However, this section of Muller discloses details of a search engine that accesses a store and forward database that controls the switch and does not disclose details of the switch itself.

Claims 5 and 48 recite that “the access interface module separates the control information and the data”. The examiner cites Muller, column 7, lines 24-32, as disclosing this limitation. This section of Muller discloses a mechanism that updates the

store and forward database based on the traffic flowing through the switch in order to “learn” on which subnet a node is located. However, the disclosed mechanism relates only to the data packets and does not disclose that control information is separated from the data as recited.

Claims 6 and 49 recite that the “data switch fabric comprises a non-blocking crossbar switch for data transfer and the control switch fabric comprises an Ethernet switch for control information transfer.” The examiner refers to Muller column 4, lines 1-23. At this section, Muller discloses details of subsystems which are used to construct a network switch. As disclosed, each subsystem includes a switch and a forwarding database. The switch is disclosed as being able to handle many formats, such as Gigabit Ethernet and Fast Ethernet. However, Muller does not disclose that the switch fabric has both a non-blocking crossbar data switch and an Ethernet control switch as claimed.

Claims 9 and 52 recite that an “access interface module selects a subset of the plurality of storage devices based on the preallocation information.” As mentioned previously, the Bachmat and Muller references perform load balancing by swapping data blocks and do not preallocate data storage locations. Therefore, these claims also distinguish over the cited reference combination. The examiner cites Bachmat, column 12, line 63 to column 13, line 8. As mentioned above at this location Bachmat discusses the sizes of blocks of data that are swapped in order to accomplish load balancing. There is no mention of preallocation of data storage locations and, as mentioned above, preallocation of data storage locations is not necessary in Bachmat/Muller because data movement occurs after the data has been stored.

Respectfully submitted

/paul e. kudirka/ Date: 2012-01-17
Paul E. Kudirka, Esq. Reg. No. 26,931
LAW OFFICES OF PAUL E. KUDIRKA
Customer Number 64967
Tel: (617) 357-0010 Email: pkudirka@paulkudirka.com

VIII APPENDIX OF CLAIMS (37 C.F.R. §41.37(c)(1)(viii))

The text of the claims involved in the appeal is:

- 1 1. Apparatus for providing high-performance, scaleable data storage services to a
2 client from a plurality of storage devices, comprising:
3 an access interface module which receives from the client data storage
4 requests, each including a data object to be stored and a data object identifier that
5 identifies that data object and, in response to each storage request and based on a
6 workload and on relative demands placed on subsets of the plurality of storage
7 devices instead of a physical location in the plurality of storage devices, dynamically
8 selects a subset of the plurality of storage devices in which the data is stored so that
9 data corresponding to the same data object identifier can be transferred to different
10 physical storage device locations from request to request in order to dynamically
11 distribute the workload across the plurality of storage devices; and
12 a switch fabric for temporarily connecting the access interface module to
13 the selected subset of the plurality of storage devices so that the data can be
14 transferred to the selected subset of storage devices.
- 1 3. The apparatus of claim 1 wherein the switch fabric comprises a control switch
2 fabric for transferring control information and a separate data switch fabric for
3 transferring data.

- 1 4. The apparatus of claim 3 wherein the control switch fabric is optimized for
2 transferring control information and the data switch fabric is optimized for
3 transferring data.
- 1 5. The apparatus of claim 3 wherein the request for storage includes control
2 information and data and wherein the access interface module separates the
3 control information and the data and transfers the data to the selected subset of
4 storage devices over the data switch fabric.
- 1 6. The apparatus of claim 3 wherein the data switch fabric comprises a non-
2 blocking crossbar switch for data transfer and the control switch fabric comprises
3 an Ethernet switch for control information transfer.
- 1 7. The apparatus of claim 1 further comprising a resource module connected to the
2 plurality of storage devices for generating preallocation information that
3 preallocates storage from the plurality of storage devices in order to evenly
4 distribute a workload across the plurality of storage devices.
- 1 8. The apparatus of claim 7 wherein the switch fabric connects the access interface
2 module to the resource module so that the resource module can transfer the
3 preallocation information to the access interface module.
- 1 9. The apparatus of claim 8 wherein the access interface module selects a subset
2 of the plurality of storage devices based on the preallocation information.

1 10. The apparatus of claim 1 wherein the access interface module comprises a data
2 memory which temporarily stores information transferred between the access
3 interface module and the selected subset of the plurality of storage devices.

1 11. The apparatus of claim 1 further comprising a plurality of access interface
2 modules each access interface module receiving storage requests from a
3 plurality of clients.

1 44. A method for providing high-performance, scalable data storage services to a
2 client from a plurality of storage devices, the method comprising:

- 3 (a) providing an access interface module which receives from the client data
4 storage requests, each including a data object to be stored and a data
5 object identifier that identifies that data object;
- 6 (b) using the access interface module in response to each data storage
7 request and based on a workload and on relative demands placed on
8 subsets of the plurality of storage devices instead of a physical location in
9 the plurality of storage devices, dynamically selects a subset of the
10 plurality of storage devices in which the data is stored so that data
11 corresponding to the same data object identifier can be transferred to
12 different physical storage device locations from request to request in order
13 to dynamically distribute the workload across the plurality of storage
14 devices; and

15 (c) using a switch fabric to temporarily connect the access interface module to
16 the selected subset of the plurality of storage devices so that the data can
17 be transferred to the selected subset of storage devices.

1 46. The method of claim 44 wherein step (c) comprises:

2 (c1) using a control switch fabric for transferring control information; and

3 (c2) using a separate data switch fabric for transferring data.

1 47. The method of claim 46 wherein step (c1) comprises optimizing the control
2 switch fabric for transferring control information and step (c2) comprises
3 optimizing the data switch fabric for transferring data.

1 48. The method of claim 46 wherein the request for storage includes control
2 information and data and wherein step (b) comprises separating the control
3 information and the data and step (c) comprises transferring the data to the
4 selected subset of storage devices over the data switch fabric.

1 49. The method of claim 46 wherein step (c1) comprises using a non-blocking
2 crossbar switch for data transfer and step (c2) comprises using an Ethernet
3 switch for control information transfer.

1 50. The method of claim 44 further comprising:

2 (d) providing a resource module connected to the plurality of storage devices;
3 and

4 (e) using the resource module to generate preallocation information that
5 preallocates storage from the plurality of storage devices in order to
6 evenly distribute a workload across the plurality of storage devices.

1 51. The method of claim 50 wherein step (c) comprises connecting the access
2 interface module to the resource module so that the resource module can
3 transfer the preallocation information to the access interface module.

1 52. The method of claim 51 wherein step (b) comprises selecting a subset of the
2 plurality of storage devices based on the preallocation information.

1 53. The method of claim 44 wherein step (b) comprises temporarily storing
2 information transferred between the access interface module and the selected
3 subset of the plurality of storage devices.

1 54. The method of claim 44 wherein step (a) further comprises providing a plurality of
2 access interface modules each access interface module receiving storage
3 requests from a plurality of clients.

IX EVIDENCE APPENDIX (37 C.F.R. §41.37(c)(1)(ix))

None

X RELATED PROCEEDINGS APPENDIX (37 C.F.R. §41.37(c)(1)(x))

No decisions have been rendered as an appeal brief has not yet been filed.